

JohanRenish / BLENDING_LERNING_Implementation-of-Multiple-Linear-Regression-Model-with-Cross-Validation

Code Pull requests Actions Projects Wiki Security Insights Settings

main / README.md

JohanRenish Update README.md adf6de6 · 1 minute ago

85 lines (65 loc) · 2.57 KB

Preview Code Blame

Raw

BLENDING_LERNING

Implementation-of-Multiple-Linear-Regression-Model-with-Cross-Validation-for-Predicting-Car-Prices

AIM:

To write a program to predict the price of cars using a multiple linear regression model and evaluate the model performance using cross-validation.

Equipments Required:

1. Hardware – PCs
2. Anaconda – Python 3.7 Installation / Jupyter notebook

Algorithm

1. Import required libraries.
2. Load and preprocess the dataset.
3. Split data into training and testing sets.
4. Train the Linear Regression model.
5. Evaluate the model using cross-validation and performance metrics.
6. Visualize actual vs predicted prices.

Program:

```
/*
Program to implement the multiple linear regression model for predicting car prices
with cross-validation.
Developed by: Johan Renish A
RegisterNumber: 212225040159
*/
import pandas as pd
from sklearn.model_selection import train_test_split,cross_val_score
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error,mean_absolute_error, r2_score
import matplotlib.pyplot as plt

#1.Load and prepare data
data=pd.read_csv('CarPrice_Assignment.csv')

# Simple preprocessing
data = data.drop(['car_ID', 'CarName'], axis=1)
data = pd.get_dummies(data, drop_first=True)

# 2.Split data
X=data.drop('price',axis=1)
y=data['price']
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=42)

#3.Create and train mode

model= LinearRegression()
model.fit(X_train,y_train)

#4. Evaluate with cross-validation(simple version)
print('Name:Johan Renish A')
print('Reg No:212225040159')
print("\n====Cross-Validation====")
cv_scores = cross_val_score(model, X, y, cv=5)
print("Fold R2 Scores:0",[f"{score:.4f}"for score in cv_scores])
print(f"Average R2:{cv_scores.mean():.4f}")

#5.Test set Evaluation

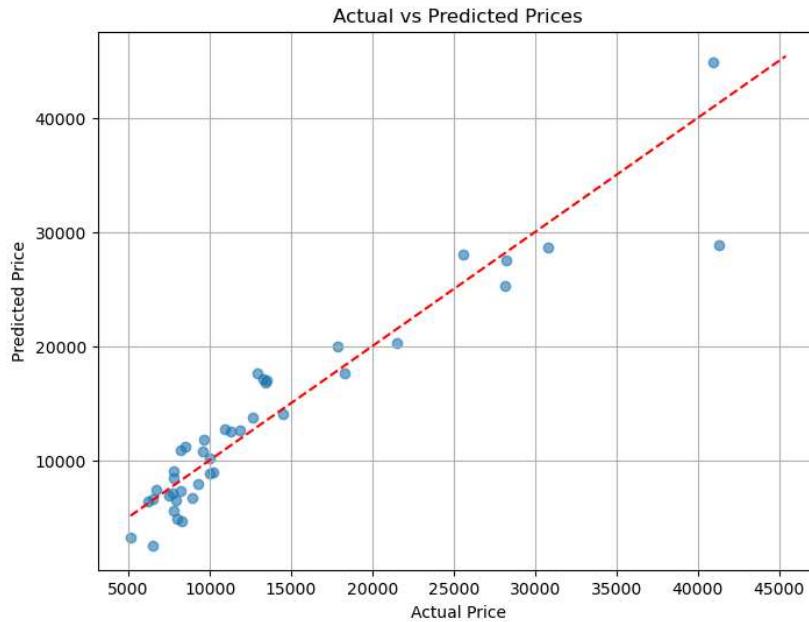
y_pred=model.predict(X_test)
print("\n==== Test Set Performance====")
print(f"MSE:{mean_squared_error(y_test,y_pred):.2f}")
print(f"MAE: {mean_absolute_error(y_test, y_pred):.2f}")
print(f"R2:{r2_score(y_test,y_pred):.4f}")

#6.Visualization
plt.figure(figsize=(8,6))
plt.scatter(y_test,y_pred, alpha=0.6)
plt.plot([y.min(), y.max()],[y.min(), y.max()], 'r--')
plt.xlabel("Actual Price")
plt.ylabel("Predicted Price")
plt.title("Actual vs Predicted Prices")
plt.grid(True)
plt.show()
```



Output:

```
Name:Johan Renish A  
Reg No:212225040159  
====Cross-Validation====  
Fold R2 Scores:0 ['0.6238', '0.6316', '0.3132', '0.3643', '-0.4944']  
Average R2:0.2877  
==== Test Set Performance====  
MSE:8482008.48  
MAE: 2089.38  
R2:0.8926
```



Result:

Thus, the program to implement the multiple linear regression model with cross-validation for predicting car prices is written and verified using Python programming.